

《心理学报》审稿意见与作者回应

题目：美在观察者眼中：陌生面孔吸引力评价中的晕轮效应和泛化效应

作者：韩尚锋 李玥 刘燊 徐强 谭群 张林

第一轮

审稿人 1 意见：

意见 1：研究将学习阶段的熟悉面孔和评价阶段的陌生面孔作为一个自变量的两个水平进行统计分析可能存在问题。因为这里所谓的陌生面孔并非真正意义上的“陌生”，而是通过软件操作有 50%的相似度，以考察泛化效应。作者仅仅比较了“100%熟悉面孔”和“50%熟悉面孔”这两种条件，缺乏与基线条件即“0%熟悉面孔”（真正的陌生面孔）的对比。因此，设置一个“完全”陌生面孔的控制条件或者设置一个不参与学习阶段的控制组，与实验条件或者实验组形成比较，才能得出该文章所作出的结论。

回应：非常感谢审稿专家的意见，我们做如下答复：

诚如审稿专家所言，将熟悉面孔和陌生面孔作为一个自变量的两个水平并不合适。我们将实验设计重新调整为单因素(特质词/行为描述的效价：积极，中性，消极)被试内设计，**修改之处见本文档 P28“2.2.3 实验设计”及 P33“3.2.3 实验设计”标蓝部分**。数据的分析将熟悉面孔的数据和陌生面孔的数据分开，按照五个步骤对数据进行分析(以实验 1 为例)：

(1)学习效果检验。对熟悉面孔的热情和能力评价进行单因素方差分析发现，对于熟悉面孔的热情评分，在不同效价水平上的得分差异显著，($F(2,43) = 43.38, p < 0.001, \eta_p^2 = 0.44$)。事后多重比较发现，与积极热情词匹配的面孔评分显著高于与中性热情词和消极热情词匹配的面孔评分，与中性热情词匹配的面孔得分显著高于与消极热情词匹配的面孔评分($ps < 0.001$)。对于熟悉面孔的能力评分，熟悉面孔在不同效价水平上的得分差异也显著，($F(2, 43) = 49.72, p < 0.001, \eta_p^2 = 0.43$)。事后多重比较发现，与积极能力词匹配的面孔评分显著高于与中性和消极能力词匹配的面孔评分，与中性能力词匹配的面孔得分显著高于与消极能力词匹配的面孔评分($ps < 0.001$)。以上结果表明面孔获得了特质词汇一致的效价，学习的效果显著。

(2)检验熟悉面孔特质对陌生面孔吸引力的影响。陌生面孔吸引力的评价差异显著，($F(2, 43) = 5.44, p < 0.01, \eta_p^2 = 0.08$)。事后多重比较发现，与积极特质面孔相似的陌生面孔吸引力评分显著高于与消极特质面孔相似的陌生面孔，与中性面孔相似的陌生面孔吸引力评分处于积极和消极面孔评价之间，与两者差异不显著($ps > 0.05$)。这表明，通过学习获得的不同面孔

特质对相似陌生面孔的吸引力评价产生了影响。

(3)检验熟悉面孔特质对熟悉面孔吸引力的影响。对熟悉面孔吸引力评分进行单因素方差分析发现，不同词汇效价匹配的熟悉面孔在吸引力评价上差异显著， $(F(2, 42) = 20.96, p < 0.001, \eta^2_p = 0.25)$ 。事后多重比较发现，与积极词匹配的面孔吸引力评分显著高于与中性词和消极词匹配的面孔吸引力评分，与中性词匹配的面孔吸引力评分也显著高于与消极词匹配的面孔吸引力评分($ps < 0.05$)。这表明熟悉面孔的吸引力评价受到面孔特质效价的影响，与积极特质匹配的面孔吸引力程度更高，即发生了晕轮效应。

(4)检验熟悉面孔特质对陌生面孔特质的影响，对陌生面孔的热情、能力评分进行单因素方差分析发现，陌生面孔的热情评分上差异显著， $(F(2, 43) = 6.19, p = 0.003, \eta^2_p = 0.09)$ 。事后多重比较发现，与积极热情面孔相似的陌生面孔评分显著高于与消极热情面孔相似的陌生面孔评分，与中性热情面孔相似的陌生面孔评分处于两者之间，与两者的差异不显著($ps > 0.05$)。这表明，陌生面孔的热情评价受到了熟悉面孔热情特质效价的影响，陌生面孔与熟悉面孔产生了效价一致的现象，即发生了热情特质的泛化。

陌生面孔的能力评分上差异显著， $(F(2, 43) = 14.18, p < 0.001, \eta^2_p = 0.18)$ ，事后多重比较发现，与积极能力面孔相似的陌生面孔在能力评分上显著高于与中性能力和消极能力面孔相似的陌生面孔评分，与中性能力面孔相似的陌生面孔能力评分显著高于与消极能力面孔相似的陌生面孔评分($ps < 0.05$)。这表明，陌生面孔的能力评价受到了熟悉面孔能力特质效价的影响，陌生面孔与熟悉面孔在能力特质上产生效价一致现象，即发生了能力特质的泛化。

(5)检验熟悉面孔特质对陌生面孔吸引力影响的路径。

表 2 各变量相关分析结果

变量	1	2	3	4	5	6	7	8	9	10	11	12
1 积极熟悉面孔特质评分	1											
2 中性熟悉面孔特质评分	0.033	1										
3 消极熟悉面孔特质评分	-0.525**	-0.118	1									
4 积极熟悉面孔吸引力评分	0.564**	-0.095	-0.452**	1								
5 中性熟悉面孔吸引力评分	0.300*	0.313*	-0.238	0.433**	1							
6 消极熟悉面孔吸引力评分	-0.384*	-0.099	0.479**	-0.082	-0.139	1						
7 积极面孔相似的陌生面孔特质评分	0.18	0.095	-0.306*	0.112	-0.135	-0.055	1					
8 中性面孔相似的陌生面孔特质评分	-0.152	0.281	0.043	-0.115	-0.082	-0.203	0.202	1				
9 消极面孔相似的陌生面孔特质评分	-0.386**	0.029	0.549**	-0.256	-0.147	0.462**	0.122	0.169	1			
10 积极面孔相似的陌生面孔吸引力评分	0.315*	-0.067	-0.415**	0.567**	0.218	-0.054	0.511**	0.096	-0.114	1		
11 中性面孔相似的陌生面孔吸引力评分	0.058	0.377*	-0.209	0.117	0.359*	-0.194	0.134	0.443**	-0.07	0.329*	1	
12 消极面孔相似的陌生面孔吸引力评分	-0.161	-0.316*	0.319*	-0.142	-0.091	0.571**	-0.004	-0.16	0.524**	-0.023	-0.2	1

注：** $p < 0.01$, * $p < 0.05$, 下同。

由表2的结果可以发现，熟悉面孔的特质评分、熟悉面孔的吸引力评分和陌生面孔的吸引力评分在积极、消极和中性三个相应水平上均有显著的相关。为了进一步考察熟悉面孔的特质评分是否通过影响其吸引力评分(发生晕轮效应)，进而影响陌生面孔的吸引力评分(出现泛化效应)，本研究采用了中介效应分析的方法进行检验。熟悉面孔的特质评价、熟悉面孔的吸引力评价和陌生面孔吸引力评价之间的相关满足中介效应检验的条件。因此，本研究采用偏差校正的 Bootstrap 法(方杰, 张敏强, 邱皓政, 2012)检验熟悉面孔吸引力评价在熟悉面孔特质评价和陌生面孔吸引力评价之间的中介作用。

检验的结果如表3所示，95%置信区间不包含0即表示中介效应成立。积极特质熟悉面孔的特质评分通过影响其吸引力评分，进而对相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平；消极特质熟悉面孔的特质评分也通过影响其吸引力评分，进而对与其相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平；中性特质的熟悉面孔则未发生相应的影响。这表明熟悉面孔的特质评价(无论是积极还是消极)均可以通过影响其吸引力评分，进而影响与其相似的陌生面孔吸引力评分，即陌生面孔的吸引力评价是由熟悉面孔特质发生晕轮效应后，进而出现吸引力的泛化效应而产生的。

消极特质的熟悉面孔特质评分与相似陌生面孔的特质评分和相似陌生面孔吸引力评分相关显著。经中介检验发现，消极特质熟悉面孔的特质评分通过影响其特质评分，进而对与其相似的陌生面孔吸引力评分产生影响的间接效应也达到显著水平。而积极特质的熟悉面孔评分与相似陌生面孔的特质评分、中性面孔特质评分与相似陌生面孔特质评分之间的相关不显著，不满足中介检验的条件。这表明熟悉面孔的消极特质评价可以通过影响其特质评分，进而影响与其相似的陌生面孔吸引力评分。

表 3 中介效应显著性检验的 Bootstrap 分析

路径	效应值	SE	Boot 95% CI
积极熟悉面孔特质评分—积极熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.16	0.13	[0.097, 0.667]
中性熟悉面孔特质评分—中性熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.13	0.10	[-0.001, 0.419]
消极熟悉面孔特质评分—消极熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.28	0.10	[0.128, 0.540]
积极熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.09	0.08	[-0.042, 0.289]
中性熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.16	0.13	[-0.011, 0.491]
消极熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.30	0.15	[0.037, 0.631]

实验 1 并未发现熟悉面孔特质评价、陌生面孔特质评价、陌生面孔吸引力评价之间在所有对应的不同水平上显著的相关，实验 2 采用更加生态化的句子材料后各变量间相关显著，

进行中介分析发现陌生面孔的特质评价在熟悉面孔特质评价和陌生面孔吸引力评价起中介作用，表明熟悉面孔特质评价也通过陌生面孔的特质评价对陌生面孔吸引力评价产生影响。**修改之处见本文档 P29“2.3.2”至 P30“2.3.6”以及 P34“3.3.2”至 P35“3.3.6”标蓝部分。**

为了能够更加严谨地证明本文的研究结论，根据审稿专家的意见，我们进行了补充实验，设置了不参与学习阶段的控制组，与实验组进行比较。我们又重新选取 35 名被试，女 19 名，男 16 名，年龄在 19~25 岁之间($M = 22.51$, $SD = 1.90$)。实验材料中选取的 12 张面孔中有 6 张经过学习成为熟悉面孔，即“100%熟悉面孔”，剩下的 6 张面孔与熟悉面孔进行 50%的合成，生成 6 张面孔为“50%熟悉面孔”，而这剩下的 6 张面孔与 6 张进行学习的面孔熟悉度为 0%，即为“0%熟悉面孔”。因此将这剩下的 6 张面孔作为基线条件让被试对其进行热情、能力和吸引力评价，结果发现：与具有积极特质熟悉面孔进行合成的完全陌生面孔热情评分为 $M = 5.40$ ，与具有中性特质熟悉面孔进行合成的完全陌生面孔热情评分为 $M = 5.00$ ，与具有消极特质熟悉面孔进行合成的完全陌生面孔热情评分为 $M = 5.27$ ，热情评分之间差异不显著($F(2, 34) = 1.58, p = 0.366$)；与具有积极特质熟悉面孔进行合成的完全陌生面孔能力评分为 $M = 5.49$ ，与具有中性特质熟悉面孔进行合成的完全陌生面孔能力评分为 $M = 5.38$ ，与具有消极特质熟悉面孔进行合成的完全陌生面孔能力评分为 $M = 5.78$ ，能力评分之间差异不显著($F(2, 34) = 1.58, p = 0.211$)；与具有积极特质熟悉面孔进行合成的完全陌生面孔吸引力评分为 $M = 4.99$ ，与具有中性特质熟悉面孔进行合成的完全陌生面孔吸引力评分为 $M = 4.86$ ，与具有消极特质熟悉面孔进行合成的完全陌生面孔吸引力评分为 $M = 5.12$ ，吸引力评分之间差异不显著($F(2, 34) = 0.39, p = 0.675$)。

该结果表明，“0%熟悉面孔”没有发生泛化，即完全陌生的面孔没有产生与熟悉面孔一致的效价。这也反过来证明与熟悉面孔相似度为 50%的陌生面孔效价的变化是由于面孔相似性所引起的情感学习泛化的作用，而不是面孔材料本身所导致的。

意见 2：另外一个需要控制的因素是模糊化处理本身是否会导致面孔吸引力的增加。作者可以找一组新的未进行实验的被试，测试他们对于“100%A”面孔，“50%A50%B”面孔和“100%B”面孔三者的吸引力评价。

回应：非常感谢审稿专家的意见。我们又重新招募被试对面孔材料进行了评定。被试一共 40 人，男 15 人，女 25 人，年龄在 19~26 岁之间($M = 22.15$, $SD = 1.99$)。实验材料中的熟悉面孔为“100%A”面孔，软件合成后的面孔为“50%A50%B”面孔，与熟悉面孔进行合成的面孔为“100%B”面孔，让被试对三种面孔的吸引力进行评价，发现三者的面孔吸引力评分差异

显著 $F(2, 39) = 4.80, p = 0.010$, 事后多重比较发现“100%A”面孔吸引力评分($M = 5.07$)与“100%B”面孔的吸引力评分($M = 4.99$)之间差异不显著, “50%A50%B”面孔吸引力评分($M = 5.64$)显著高于“100%A”和“100%B”面孔吸引力评分($ps < 0.38$)。该结果表明合成后面孔的吸引力确实有所增加, 与以往的研究结果一致(Langlois & Roggman, 1990; Trujillo, Jankowitsch, & Langlois, 2014)。

但是“100%A”面孔中与积极、中性和消极信息匹配的面孔吸引力评分之间差异不显著($F(2, 39) = 0.35, p = 0.706$); “100%B”面孔中与积极、中性和消极信息匹配的面孔吸引力评分之间差异不显著($F(2, 39) = 0.39, p = 0.675$); “50%A50%B”面孔中与积极、中性和消极信息匹配的面孔吸引力评分之间差异也不显著($F(2, 39) = 1.97, p = 0.143$)。

虽然合成后的面孔吸引力有所提升, 但是熟悉面孔不同效价之间、合成面孔不同效价之间以及完全陌生面孔不同效价之间没有显著差异, 这表明正式实验中熟悉面孔吸引力评价的差异是由于与不同效价的信息进行匹配学习所导致, 陌生面孔吸引力评价的差异是晕轮效应和泛化效应所导致。因此, 面孔处理本身并不是熟悉面孔和陌生面孔吸引力评价发生差异的原因。

涉及的相关文献:

- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*(2), 115-121.
- Trujillo, L. T., Jankowitsch, J. M., & Langlois, J. H. (2014). Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cognitive Affective & Behavioral Neuroscience, 14*(3), 1061-1076.

意见 3: 文章写作的逻辑性需进一步加强。根据实验程序可知, 实验的主要逻辑是特质词与面孔匹配产生晕轮效应, 这种效应会通过面孔的相似性传递或者泛化。但是在写作的时候, 作者并未按照这一逻辑将二者联系起来, 而是将情感学习泛化和晕轮效应作为两个单独的过程提出了两个独立的假设。这一点在讨论部分也需注意, 需要基于上述实验逻辑进一步结合这两个机制来讨论, 而不只是分开讨论。

回应: 非常感谢审稿专家具有建设性的意见。我们经过文献梳理, 发现陌生面孔吸引力的形成过程中, 晕轮效应后发生泛化效应是陌生面孔吸引力形成的一种途径, 而且在实验 1 和实验 2 中我们通过增加中介分析发现, 熟悉面孔的吸引力在熟悉面孔特质和陌生面孔吸引力之间起着中介作用, 在知觉陌生面孔吸引力时, 熟悉面孔的特质首先影响了熟悉面孔的吸引力, 即发生晕轮效应, 之后熟悉面孔的吸引力发生泛化影响了陌生面孔吸引力, 即发生泛化效应。

然而我们之前提出的熟悉面孔印象发生泛化进而产生晕轮效应这条路径也是存在的,尽管实验 1 并未作出陌生面孔特质在熟悉面孔特质和陌生面孔吸引力之间的中介作用,但是实验 2 采用生态效度更高的句子材料后发现了陌生面孔特质在熟悉面孔特质和陌生面孔吸引力之间的中介作用。因此熟悉面孔的特质发生泛化进而引起晕轮效应这条途径也是存在的。

我们结合了这两条途径对文章进行了修改,在正文第五段先对晕轮效应进行了论述,第六段对泛化效应进行了论述,第七段进一步说明泛化效应发生以面孔相似性为基础,第八段结合晕轮效应和泛化效应提出假设陌生面孔吸引力的形成可能存在的两条途径。**修改之处见本文档 P25 最后一段至 P27 第二段。**

具体修改为:“个体在形成对他人印象的过程中会发生晕轮效应(halo effect)。Gross 和 Crofton(1977)发现,无论面孔是高吸引力、中等吸引力还是低吸引力,和受人喜爱、中等、不受喜爱的描述匹配后,在吸引力的三个水平上和积极描述相匹配的面孔吸引力评价都更高。Zhang, Kong, Zhong 和 Kou(2014)的研究也发现,被试对不同效价的特质词和面孔分别进行匹配学习后,会认为积极词匹配的面孔比中性和消极词匹配的面孔更有吸引力。个体对他人的印象越积极,就觉得对方越有吸引力,会导致“好既是美”(what-is-good-is-beautiful effect)的现象,即发生了晕轮效应(Zebrowitz & Franklin, 2014)。刻板印象内容模型(stereotype content model)提出社会认知内容主要包括热情和能力两个维度,热情体现在与人交往的过程中,是为了维持良好的社会关系,包括真诚、友好、信任等品质;而能力则体现在追求目标的过程中,能够使特质拥有者获益,包括聪明、独立、工作高效等品质(Abele & Bruckmüller, 2011; Abele & Wojciszke, 2014)。以往的研究表明,热情特质会对面孔吸引力产生积极影响,表现为与积极热情特质匹配的面孔比消极热情特质匹配的面孔会更有吸引力。例如,相比于具有不可信特质的男性面孔,有可信特质的男性面孔更容易被女性知觉为具有更高的吸引力(Quist, DeBruine, Little & Jones, 2012);能力特质也会对面孔吸引力评价产生影响,例如王雨晴等人(2015)的研究发现具有积极能力特质的面孔更有吸引力。因此,面孔的热情和能力特质都可能通过晕轮效应对陌生面孔的吸引力评价产生影响。

泛化效应(generalization effect)的发生是陌生面孔印象形成中一个重要的过程(Kocsor & Bereczkei, 2017; 张林, 李玥, 刘燊, 谭群, 徐强, 杨亚平, 2016)。Verosky和Todorov(2010)发现通过情感学习,面孔的印象会发生泛化。他们让被试对积极、中性、消极的行为描述和面孔进行匹配学习获得面孔的社会信息,然后对与其相似的陌生面孔进行特质评价,结果发现当熟悉面孔具有积极特质时,与其相似的陌生面孔评价也是积极的,而熟悉面孔具有消极特质时,与其相似的陌生面孔评价也更消极,即发生了情感学习泛化(affective learning generalization)。情感学习泛化是

陌生面孔和熟悉面孔由于相似而产生一致印象的过程。之后有研究发现，不仅具有社会属性的情感信息与面孔匹配学习会发生泛化，非社会属性的情感信息与面孔进行学习形成联结后，面孔获得的情感效价也会发生泛化。Kocsor和Bereczkei (2017)发现积极的情感图片与面孔学习形成联结后，被试对与其相似的陌生面孔评价也更积极。此外，泛化不仅会使个体间形成一致的印象，也会使个体与群体、群体与群体形成一致的印象(Glaser, Dickel, Liersch, Rees, Sussenbach, & Bohner, 2015)。因此，在陌生面孔吸引力印象形成过程中，熟悉面孔的泛化效应可能也会发挥重要作用。

泛化效应的发生需要基于面孔特征间的相似性，个体通过对面孔相似性的加工从而形成对他人不同的印象和评价(李玥, 尹月阳, 张林, 2015; 尹月阳, 张林, 李玥, 杨亚平, 2015)。当看到陌生人的面孔与熟人的面孔相似时，个体就会认为陌生人和熟人具有相同的特质(Kraus & Chen, 2010)。求职者面孔与高能力的员工相似时会得到更高的评价，与低能力的员工相似时会得到更低的评价(Von, Herzog, & Rieskamp, 2014)。有研究也发现，人们在对与自我相似的面孔进行判断时会有更积极的评价(Richter, Tiddeman, & Haun, 2016)。人们利用相似性信息对面孔进行加工时具有自动化的特点(Gawronski & Quinn, 2013; Verosky & Todorov, 2013)，这种自动化的加工能够使观察者通过面孔相似性对陌生人快速形成第一印象。同时，张林等人(2016)也指出，对陌生面孔和熟悉面孔相似线索的加工是形成陌生面孔印象的一种重要方式。因此，知觉到陌生面孔与熟悉面孔的相似性线索可能是影响面孔印象发生泛化的重要前提。

先前的研究发现泛化效应存在于同种属性之间，如 Verosky 和 Todorov(2010, 2013)的研究中发现熟悉面孔的信任印象会泛化到陌生面孔，Zebrowitz, White 和 Wieneke(2008)对熟悉面孔的喜爱度也会发生泛化而影响对陌生面孔的喜爱度。基于面孔晕轮效应和泛化效应两种加工过程，本文提出个体在对陌生面孔吸引力进行加工时，不仅熟悉面孔的特质会发生晕轮效应和泛化效应，陌生面孔特质也会发生泛化和晕轮效应。一方面，熟悉面孔的积极特质会出现晕轮效应导致其面孔吸引力提高，进而通过泛化影响陌生面孔吸引力的加工，即熟悉面孔特质出现晕轮效应后发生泛化效应可能是陌生面孔吸引力加工的途径之一；另一方面，面孔特质的泛化效应也会使陌生面孔与熟悉面孔特质具有一致的效价，具有积极特质的熟悉面孔会通过泛化使其相似的陌生面孔也具有相同效价的特质，进而发生晕轮效应进一步提高其吸引力评价，即熟悉面孔特质发生泛化后，进而出现晕轮效应可能是陌生面孔吸引力加工的另一条途径。”

在讨论部分按照陌生面孔吸引力发生的两条途径，即晕轮效应后发生泛化效应，泛化效应后发生晕轮效应这两种途径进行讨论，**修改之处见本文档 P38 第二段和第三段。**

具体修改为：“陌生面孔吸引力知觉过程包括两种途径，晕轮效应后发生泛化效应是陌

生面孔吸引力加工的途径之一。有研究发现对于熟悉面孔会发生晕轮效应，表现为那些具有积极特质的面孔，其吸引力评价也更高。而情感学习泛化会使陌生面孔和相似的熟悉面孔具有一致的效价(Gawronski & Quinn, 2013; Verosky & Todorov, 2010, 2013)，本研究也得到相同的结论。晕轮效应后发生泛化效应可以解释在我们生活中虽然与某些人素未谋面，但却产生了莫名的好感或者厌恶。这种感觉可能是由于人们看到的陌生面孔与之前熟人的面孔相似，熟人的印象泛化到了陌生人的面孔上。此外，本研究发现对于具有不同热情和能力效价的熟悉面孔，不仅热情和能力特质发生了泛化，而且吸引力也发生了泛化。因此，情感学习的泛化不仅只局限于相同特质的泛化，相近或相似的特质也会发生泛化。本研究结果表明除了面孔的热情和能力特质可以发生泛化之外，吸引力特质也发生了泛化现象，这也表明泛化效应存在的普遍性。虽然泛化效应在陌生面孔加工中具有重要作用，泛化如何发生还不清楚。本研究发现面孔相似度为 50%的情况下发生了泛化，与前人的结果一致 (Verosky & Todorov, 2013)，说明相似性是情感学习泛化效应发生的必要条件。然而，Kocsor 和 Bereczkei(2016)通过让被试根据行为描述使面孔形成可信和不可信的印象，结果发现被试没有觉察到相似性的情况下，对积极行为匹配的面孔评价也更为可信。他们认为这种泛化是神经系统普遍的功能，泛化效应是熟悉面孔行为特质的泛化，与面孔相似性的觉察与否无关。因此，未来对泛化效应的内在机制还有待于进一步探讨。

泛化效应导致晕轮效应是陌生面孔吸引力知觉加工的另一个途径。本研究发现，对陌生面孔进行知觉时也会发生晕轮效应。陌生面孔吸引力知觉是如何发生的，还有不同解释。例如常模参照模型(norm-based coding model, NBC)认为，人们在记忆中储存着很多常模脸或平均脸，面孔刺激与常模脸或平均脸进行对比，根据偏差识别不同的面孔，而面孔的不同属性(如性别、种族)有不同常模(Valentine, 1988)。因此，在记忆中人们可能储存着高吸引力的常模脸，在对面孔吸引力进行评价时知觉到的面孔和高吸引力常模脸越相似，越有可能被识别为高吸引力面孔。该模型有两个局限：第一，该模型不能解释为何那些具有积极特质的面孔也可以被知觉为有更高的吸引力；第二，记忆中如果储存了大量常模脸会增加记忆负担，导致匹配过程变得缓慢，而面孔吸引力判断是自动化的过程。晕轮效应的发生意味着个体只需要知觉到积极或消极效价就可以产生不同的吸引力判断，既可以减轻记忆的负担，又可以迅速形成吸引力知觉。因此，通过熟悉面孔的泛化进而发生晕轮效应这种方式更加高效，可能是陌生面孔吸引力快速判断的原因。”

意见 4：文章题目为“一见钟情：陌生面孔吸引力的加工机制”，题目可能有点以偏概全，并不能精准地表述文章的主题。文章主要考察晕轮效应是否会通过面孔的相似性泛化，这可能

只是一见钟情的潜在机制之一。

回应：非常感谢审稿专家的意见。我们经过仔细思考后发现，原来的题目不能够准确地表达文章的主题，本文主要探讨陌生面孔吸引力形成的加工过程，强调在形成过程中晕轮效应和泛化效应共同发挥的作用，因此我们将题目修改为“陌生面孔吸引力的加工:晕轮效应和泛化效应”，英文的题目也进行了修改并用蓝色字体标出，**修改之处见本文档 P24 正文标题及 P43 英文标题**，

.....

审稿人 2 意见：

意见 1：某些专业的名词，作者应给出相应的英文原文，如 P3 中，面孔所有者假设。P4 中，面孔的观察者假设。请作者检查全文，对专业名词给出相应英文原文，方便其他研究者阅读。

回应：非常感谢审稿专家的意见。我们对全文中的专业名词进行了核查，补充了相应的英文原文，具体修改情况为:面孔所有者假设(the owner hypothesis)、面孔的观察者假设(the observer hypothesis)，**修改之处分别见本文档 P24 最后一段第一行、P25 页第二段第二行**。

另外，我们对文章进行了仔细的检查，对文中其他的专业名词进行了补充和修正，如荷尔蒙(hormone)、常模参照模型(norm-based coding model, NBC)，**修改之处分别见本文档 P25 第二段第五行、P38 最后一段第三行**，已用蓝色字体标出。

意见 2：在实验 1 的结果 2.3.2 中，对面孔热情度的评价。对熟悉面孔的评价，是为了检验学习效果，对陌生面孔的评价，是为了检验熟悉面孔对陌生面孔的影响。将熟悉面孔与陌生面孔的数据放在一起进行方差分析是否合适？

例如，词汇效价主效应显著，作者得出的结论是“积极词匹配的面孔热情评分显著大于中性词和消极词匹配的面孔热情评分,中性词匹配的面孔热情评分也显著大于消极词匹配的面孔热情评分”，而这种结论并不合适，因为词汇效价的主效应是结合了熟悉面孔与陌生面孔的数据的检验结果，但只有熟悉面孔与不同效价的词汇进行了匹配，陌生面孔并没有与不同效价的词汇进行匹配，因此不能得出上述的结论。

此外，交互效应显著，进而进行简单分析，发现不同词汇效价匹配的熟悉面孔热情评价差异显著，然后发现陌生面孔的热情评价差异也显著，进而得出“陌生面孔的热情评价受到熟悉面孔特质效价的影响,陌生面孔与熟悉面孔产生效价一致的热情印象,即发生了泛化效应。”这样的结论，这种推理合适吗？

是否应该将熟悉面孔与陌生面孔的数据分开分析，首先分析熟悉面孔的效应，词汇效

价的效应显著，表明学习效果好。然后分析陌生面孔的数据，词汇效价的效应也显著，同时与对应的熟悉面孔之间存在高相关。进而得出“陌生面孔的热情评价受到熟悉面孔特质效价的影响”这样的结论。

对面孔能力评价的数据分析存在的问题同上？

回应：感谢审稿专家的宝贵意见和建议，我们作如下答复：

(1)诚如审稿专家所言，将熟悉面孔与陌生面孔数据放在一起分析并不合适。也难以得出“积极词匹配的面孔热情评分显著大于中性词和消极词匹配的面孔热情评分,中性词匹配的面孔热情评分也显著大于消极词匹配的面孔热情评分”的结论。

(2)我们查阅了以往文献中对泛化效应的解释以及分析方法。

根据 Verosky 和 Todorov(2010, 2013)的研究，泛化是指与积极面孔相似的面孔评价比消极面孔相似的面孔评价更积极，原文为：“Faces that resembled positive faces were evaluated more positively than faces that resembled negative faces.”因此，泛化效应的发生需要证明与积极面孔相似的陌生面孔比与消极面孔相似的陌生面孔更积极。

具体的分析方法是：首先证明学习效果显著，即与不同效价词汇或句子进行匹配后中性的面孔效价发生了变化，与消极信息匹配的面孔比与中性和积极信息匹配的面孔更消极，与中性信息匹配的面孔比积极信息匹配的面孔更消极。

Verosky 和 Todorov(2010)的研究中证明学习效果的方法是：Analysis of the judgments of the learned faces revealed a large effect of the valence of the behavior, $F(2,112) = 93.57, p < 0.001, \eta^2_p = .63$. Participants evaluated faces that were previously associated with negative behaviors ($M = 3.02, SD = 1.35$) more negatively than faces that were associated with neutral behaviors ($M = 5.50, SD = 1.15$), $t(56) = 10.90, p < 0.001$. In turn, they rated the latter more negatively than faces that were associated with positive behaviors ($M = 6.56, SD = 1.59$), $t(56) = 4.54, p < .001$.

Verosky 和 Todorov(2013)的研究中证明学习效果的方法是：The biographical information influenced participants' evaluation of the familiar faces, $F(2, 140) = 28.09, p < .001, \eta^2_p = .29$, with participants evaluating the face paired with the negative biography ($M = 3.63, SD = 2.36$) significantly more negatively than the face paired with the neutral ($M = 5.68, SD = 1.60; t(70) = 6.53, p < .001$) or positive biography ($M = 6.25, SD = 2.18; t(70) = 5.76, p < .001$). Participants also evaluated the face paired with the neutral biography more negatively than the face paired with the positive biography, though this difference was only marginally significant, $t(70) = 1.83, p = .07$.

然后证明情感学习的泛化,与具有积极效价熟悉面孔相似的面孔评价比与中性和消极效价熟悉面孔相似的陌生面孔更积极,与具有中性效价熟悉面孔相似的陌生面孔评价比与消极效价熟悉面孔相似的陌生面孔更积极。

Verosky 和 Todorov(2010)的研究中证明发生情感学习泛化的方法是“participants rated novel faces morphed with positive learned faces ($M = 5.25, SD = 0.94$) more positively than novel faces morphed with neutral ($M = 5.20, SD = 0.87$) or negative ($M = 4.99, SD = 0.90$) faces, $F(2, 106) = 3.09, p < .05, \eta_p^2 = .06$.”

Verosky 和 Todorov(2013)的研究中证明情感学习泛化的方法是: Importantly, learning generalization also occurred, as evidenced by a significant effect of valence of learning, $F(2, 140) = 6.63, p = .002, \eta_p^2 = .09$. Participants evaluated faces that were similar to the face paired with the negative biography ($M = 4.98, SD = 1.06$) more negatively than those that were similar to the face paired with the neutral ($M = 5.26, SD = .94; t(70) = 2.89, p = .005$) or positive biography ($M = 5.31, SD = 1.04; t(70) = 3.06, p = .003$). Participants also evaluated faces that were similar to the neutral face more negatively than those that were similar to the positive face, though this difference was not significant.

我们根据前人研究的分析方法以及专家的建议,对数据的分析重新进行了调整。

对学习效果的检验只分析了熟悉面孔的数据。将熟悉面孔按照与词汇匹配的效价分为积极、中性、消极三个组,进行单因素方差分析,发现学习后面孔效价与词汇效价一致,表明学习效果显著。原文修改为:“对熟悉面孔的热情和能力评价进行单因素方差分析发现,对于熟悉面孔的热情评分,在不同效价水平上的得分差异显著, ($F(2,43) = 43.38, p < 0.001, \eta_p^2 = 0.44$)。事后多重比较发现,与积极热情词匹配的面孔评分显著高于与中性热情词和消极热情词匹配的面孔评分,与中性热情词匹配的面孔得分显著高于与消极热情词匹配的面孔评分 ($ps < 0.001$)。对于熟悉面孔的能力评分,熟悉面孔在不同效价水平上的得分差异也显著, ($F(2, 43) = 49.72, p < 0.001, \eta_p^2 = 0.43$)。事后多重比较发现,与积极能力词匹配的面孔评分显著高于与中性和消极能力词匹配的面孔评分,与中性能力词匹配的面孔得分显著高于与消极能力词匹配的面孔评分 ($ps < 0.001$)。以上结果表明面孔获得了特质词汇一致的效价,学习的效果显著。”

然后分析泛化效应,陌生面孔按照与相似熟悉面孔的效价分为积极、中性、消极三个组,对于陌生面孔的热情评分,发现与积极熟悉面孔相似的陌生面孔比与消极熟悉面孔相似的陌生面孔更积极,与中性熟悉面孔相似的陌生面孔处于前两者之间。因此与积极效价熟悉面孔

相似的陌生面孔评价也更积极，与消极效价熟悉面孔相似的陌生面孔也更消极，陌生面孔与熟悉面孔获得一致的效价。该结果表明陌生面孔的效价受到了熟悉面孔特质效价的影响，即发生了泛化效应。原文修改为：“对陌生面孔的热情、能力评分进行单因素方差分析发现，陌生面孔的热情评分上差异显著， $(F(2, 43) = 6.19, p = 0.003, \eta_p^2 = 0.09)$ 。事后多重比较发现，与积极热情面孔相似的陌生面孔评分显著高于与消极热情面孔相似的陌生面孔评分，与中性热情面孔相似的陌生面孔评分处于两者之间，与两者的差异不显著($ps > 0.05$)。这表明，陌生面孔的热情评价受到熟悉面孔热情特质效价的影响，陌生面孔与熟悉面孔产生了效价一致的现象，即发生了热情特质的泛化。

陌生面孔的能力评分上差异显著， $(F(2, 43) = 14.18, p < 0.001, \eta_p^2 = 0.18)$ ，事后多重比较发现，与积极能力面孔相似的陌生面孔在能力评分上显著高于与中性能力和消极能力面孔相似的陌生面孔评分，与中性能力面孔相似的陌生面孔能力评分显著高于与消极能力面孔相似的陌生面孔评分($ps < 0.05$)。这表明，陌生面孔的能力评价受到了熟悉面孔能力特质效价的影响，陌生面孔与熟悉面孔在能力特质上产生效价一致现象，即发生了能力特质的泛化。”

我们也根据专家的意见做了熟悉面孔特质与陌生面孔特质的相关，尽管实验 1 并未出现熟悉面孔特质与陌生面孔特质在对应的不同效价水平上显著的相关，具有积极和中性特质熟悉面孔评分和与其相似的陌生面孔特质评分之间相关系数较低，相关系数分别 0.180 和 0.281，但实验 2 采用更具生态化的句子材料发现了熟悉面孔特质与陌生面孔特质在对应的不同效价水平上显著的相关，相关系数分别为 0.477、0.408、0.540。

同样，我们对面孔能力评价数据分析方法也进行了调整。**修改之处见本文档 P30“2.3.5”、“2.3.6”以及 P35“3.3.5”、“3.3.6”。**

涉及的相关文献：

Verosky, S. C., & Todorov, A. (2010). Generalization of affective learning about faces to perceptually similar faces. *Psychological Science, 21*(6), 779-785.

Verosky, S. C., & Todorov, A. (2013). When physical similarity matters: Mechanisms underlying affective learning generalization to the evaluation of novel faces. *Journal of Experimental Social Psychology, 49*(4), 661-669.

意见 3：在实验 1 的结果 2.3.2 中，对面孔吸引力的评价。首先词汇效价的主效应，存在同问题 2 中的问题。其次，交互效应显著后，进行简单效应分析，词汇效价对熟悉面孔吸引力的影响显著，进而得出晕轮效应，这没有问题，但词汇效价对陌生面孔吸引力的影响也显著，作者得出“也发生了面孔特质的晕轮效应”，这一结论存在问题，因为陌生面孔本身并没有直

接与不同效应的词汇进行匹配过。这实际上是熟悉面孔的晕轮效应“迁移”到了陌生面孔上，是晕轮效应的泛化，而非陌生面孔直接发生了晕轮效应。那么在讨论中 P14，作者是否还能得出“晕轮效应是陌生面孔吸引力形成的另一个加工过程。”这样的结论呢？因为陌生面孔本身并没有直接发生晕轮效应，而是熟悉面孔的晕轮效应泛化到陌生面孔上。请作者仔细思考。

回应：感谢审稿专家的宝贵意见。如前所述，我们根据已有的文献和专家的意见对文章的逻辑重新进行了梳理。我们根据专家的意见将熟悉面孔和陌生面孔的数据分开，按照以下步骤对数据进行分析，原文修改情况可参见意见 6：

(1)对熟悉面孔特质进行分析发现面孔获得了与词汇一致的效价，表明学习效果显著。

(2)分析陌生面孔吸引力的数据，与不同特质效价熟悉面孔相似的陌生面孔吸引力评价之间差异显著，得出熟悉面孔的特质影响陌生面孔吸引力。

(3)对熟悉面孔的吸引力进行分析，发现熟悉的面孔越积极，吸引力评价越高，即发生晕轮效应。

(4)对陌生面孔特质进行分析，结果表明与积极特质面孔相似的陌生面孔比与消极特质面孔相似的陌生面孔更积极，即发生了泛化效应。

(5)通过相关分析发现各变量间具有高相关，并进行中介效应检验发现，在实验 1 中发现熟悉面孔特质、熟悉面孔吸引力和陌生面孔吸引力对应的不同水平之间相关显著，熟悉面孔吸引力在熟悉面孔特质和陌生面孔吸引力之间起着中介作用，表明熟悉面孔特质首先影响熟悉面孔吸引力发生晕轮效应，熟悉面孔吸引力又影响到陌生面孔吸引力，发生了泛化效应。因此熟悉面孔特质通过熟悉面孔吸引力影响了陌生面孔吸引力。因此熟悉面孔特质发生晕轮效应之后产生泛化效应是陌生面孔吸引力形成的一种途径。

由于实验材料的原因，熟悉面孔特质、陌生面孔特质和陌生面孔吸引力之间对应的不同水平之间的相关不完全显著。但在实验 2 中采用生态化效度更高的句子材料后，熟悉面孔特质、陌生面孔特质和陌生面孔吸引力之间对应的不同水平之间的相关显著。通过中介分析发现陌生面孔的特质在熟悉面孔特质与陌生面孔吸引力之间起着中介作用，熟悉面孔特质对陌生面孔特质产生影响，即发生了泛化效应，陌生面孔特质对陌生面孔吸引力产生影响，即发生了晕轮效应。熟悉面孔特质也通过陌生面孔特质对陌生面孔吸引力产生影响，在此过程中陌生面孔也发生了晕轮效应。

因此，熟悉面孔特质对陌生面孔的影响有两条途径，一种是熟悉面孔发生晕轮效应后产生泛化效应，另一种是熟悉面孔特质发生泛化进而导致晕轮效应。这两种途径中晕轮效应和泛化效应共同发生作用，都是陌生面孔吸引力形成的加工过程。

意见 4: 在实验 1 的结果中, 无论是热情、能力、吸引力, 均出现了面孔类型的主效应, 即陌生面孔的热情、能力、吸引力均高于熟悉面孔, 而作者对这三个数据检验结果, 均做了回避, 没有进行任何解释。这些结果与现有的部分文献存在一定争议, 如现有的文献, 有些文献表明对熟悉面孔的评价高于对陌生面孔的评价。请作者对自己的研究结果给出合理的解释。

回应: 非常感谢审稿专家的意见, 我们作如下答复:

陌生面孔的热情、能力、吸引力均高于熟悉面孔可能是由于处理技术导致的, 由于这些陌生面孔是通过电脑软件合成得到的, 已有的研究表明合成后的面孔会更有吸引力(Langlois & Roggman, 1990; Trujillo, Jankowitsch, & Langlois, 2014)。而且, 审稿人 1 意见 2 中我们进行补充实验发现, 合成后的面孔吸引力有所增加, 这种面孔的处理可能是陌生面孔吸引力高于熟悉面孔吸引力的原因。

有研究也发现“美既是好”的现象, 高吸引力的面孔也具有更加积极的品质(Dion, Berscheid, & Walster, 1972; Tsukiura & Cabeza, 2011; Zhang, Zheng, & Wang, 2016)。因此导致陌生面孔的热情、能力评分均高于熟悉面孔的热情、能力评分。

面孔合成技术只是导致陌生面孔比熟悉面孔热情、能力、吸引力更高, 并不会导致陌生面孔的效价产生差异。因此陌生面孔效价的变化是通过匹配学习发生学习泛化导致的。

涉及的相关文献:

Dion, K. K., Berscheid, E., & Walster, E. (1972). WHAT IS BEAUTIFUL IS GOOD. *Journal of Personality and Social Psychology*, 24(3), 285-290.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1(2), 115-121.

Trujillo, L. T., Jankowitsch, J. M., & Langlois, J. H. (2014). Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cognitive Affective & Behavioral Neuroscience*, 14(3), 1061-1076.

Tsukiura, T., & Cabeza, R. (2011). Shared brain activity for aesthetic and moral judgments: implications for the Beauty-is-Good stereotype. *Social Cognitive and Affective Neuroscience*, 6(1), 138-148.

Zhang, Y., Zheng, M., & Wang, X. (2016). Effects of facial attractiveness on personality stimuli in an implicit priming task: an ERP study. *Neurological Research*, 38(8), 685-691.

意见 5: 作者在全文省略了非常非常多的“与”字, 导致语句不通顺, 请通读检查。如 P9: 与

积极面孔相似的陌生面孔能力评分显著高于“与”中性面孔相似的陌生面孔,“与”中性面孔相似的陌生面孔能力评分高于“与”消极面孔相似的陌生面孔。又如 P9: 表现为在知觉那些与积极特质熟悉面孔相似的陌生面孔时评价更积极;在知觉那些“与”消极特质熟悉面孔相似的陌生面孔时更消极。

回应: 非常感谢审稿专家的细致严谨的审阅并耐心指出文中的不足。我们对文中的句子进行了通读检查,对文章书写欠妥之处进行了修改,以保证语句准确流畅。

原来 P9“与积极面孔相似的陌生面孔能力评分显著高于‘与’中性面孔相似的陌生面孔,‘与’中性面孔相似的陌生面孔能力评分高于‘与’消极面孔相似的陌生面孔”根据本次修改内容已将句子改为“与积极能力面孔相似的陌生面孔能力评分显著高于与中性能力和消极能力面孔相似的陌生面孔评分,与中性能力面孔相似的陌生面孔能力评分显著高于与消极能力面孔相似的陌生面孔评分”,修改之处见本文档 P30 最后一段第二行。

由于我们对文章内容进行了更改,原来 P9“表现为在知觉那些与积极特质熟悉面孔相似的陌生面孔时评价更积极;在知觉那些‘与’消极特质熟悉面孔相似的陌生面孔时更消极”被删除。

此外,我们对文中其他语句不通的地方进行了修改,已在文中用蓝色字体标出,因修改之处较多,在此不便一一列举。再次感谢审稿专家耐心细致的审阅。

意见 6: 实验 2 的结果分析存在同实验 1 的结果分析同样的问题,请作者仔细思考数据分析方法的合理性,如果作者坚持自己数据分析方法的合理性,请给出能说服我的理由,如果作者经过思考,觉得确实存在问题,请选择更为合适的方式进行数据分析。

回应: 非常感谢审稿专家的意见。我们查阅了相关的文献,根据前人的分析方法以及审稿专家的意见,实验 1 的数据分析修改如下:

2.3.2 学习效果检验

对熟悉面孔的热情和能力评价进行单因素方差分析发现,对于熟悉面孔的热情评分,在不同效价水平上的得分差异显著, ($F(2,43) = 43.38, p < 0.001, \eta^2_p = 0.44$)。事后多重比较发现,与积极热情词匹配的面孔评分显著高于与中性热情词和消极热情词匹配的面孔评分,与中性热情词匹配的面孔得分显著高于与消极热情词匹配的面孔评分($ps < 0.001$)。对于熟悉面孔的能力评分,熟悉面孔在不同效价水平上的得分差异也显著, ($F(2, 43) = 49.72, p < 0.001, \eta^2_p = 0.43$)。事后多重比较发现,与积极能力词匹配的面孔评分显著高于与中性和消极能力词匹配的面孔评分,与中性能力词匹配的面孔得分显著高于与消极能力词匹配的面孔评分($ps <$

0.001)。以上结果表明面孔获得了特质词汇一致的效价，学习的效果显著。

2.3.3 熟悉面孔特质对陌生面孔吸引力的影响

陌生面孔吸引力的评价差异显著，($F(2, 43) = 5.44, p < 0.01, \eta_p^2 = 0.08$)。事后多重比较发现，与积极特质面孔相似的陌生面孔吸引力评分显著高于与消极特质面孔相似的陌生面孔，与中性面孔相似的陌生面孔吸引力评分处于积极和消极面孔评价之间，与两者差异不显著($ps > 0.05$)。这表明，通过学习获得的不同面孔特质对相似陌生面孔的吸引力评价产生了影响。

2.3.4 熟悉面孔的特质对熟悉面孔吸引力的影响

对熟悉面孔吸引力评分进行单因素方差分析发现，不同词汇效价匹配的熟悉面孔在吸引力评价上差异显著，($F(2, 42) = 20.96, p < 0.001, \eta_p^2 = 0.25$)。事后多重比较发现，与积极词匹配的面孔吸引力评分显著高于与中性词和消极词匹配的面孔吸引力评分，与中性词匹配的面孔吸引力评分也显著高于与消极词匹配的面孔吸引力评分($ps < 0.05$)。这表明熟悉面孔的吸引力评价受到面孔特质效价的影响，与积极特质匹配的面孔吸引力程度更高，即发生了晕轮效应。

2.3.5 熟悉面孔的特质对陌生面孔特质的影响

对陌生面孔的热情、能力评分进行单因素方差分析发现，陌生面孔的热情评分上差异显著，($F(2, 43) = 6.19, p = 0.003, \eta_p^2 = 0.09$)。事后多重比较发现，与积极热情面孔相似的陌生面孔评分显著高于与消极热情面孔相似的陌生面孔评分，与中性热情面孔相似的陌生面孔评分处于两者之间，与两者的差异不显著($ps > 0.05$)。这表明，陌生面孔的热情评价受到熟悉面孔热情特质效价的影响，陌生面孔与熟悉面孔产生了效价一致的现象，即发生了热情特质的泛化。

陌生面孔的能力评分上差异显著，($F(2, 43) = 14.18, p < 0.001, \eta_p^2 = 0.18$)，事后多重比较发现，与积极能力面孔相似的陌生面孔在能力评分上显著高于与中性能力和消极能力面孔相似的陌生面孔评分，与中性能力面孔相似的陌生面孔能力评分显著高于与消极能力面孔相似的陌生面孔评分($ps < 0.05$)。这表明，陌生面孔的能力评价受到了熟悉面孔能力特质效价的影响，陌生面孔与熟悉面孔在能力特质上产生效价一致现象，即发生了能力特质的泛化。

2.3.6 熟悉面孔特质对陌生面孔吸引力影响的路径分析

表 2 各变量相关分析结果

变量	1	2	3	4	5	6	7	8	9	10	11	12
1 积极熟悉面孔特质评分	1											
2 中性熟悉面孔特质评分	0.033	1										
3 消极熟悉面孔特质评分	-0.525**	-0.118	1									
4 积极熟悉面孔吸引力评分	0.564**	-0.095	-0.452**	1								
5 中性熟悉面孔吸引力评分	0.300*	0.313*	-0.238	0.433**	1							
6 消极熟悉面孔吸引力评分	-0.384*	-0.099	0.479**	-0.082	-0.139	1						
7 积极面孔相似的陌生面孔特质评分	0.18	0.095	-0.306*	0.112	-0.135	-0.055	1					
8 中性面孔相似的陌生面孔特质评分	-0.152	0.281	0.043	-0.115	-0.082	-0.203	0.202	1				
9 消极面孔相似的陌生面孔特质评分	-0.386**	0.029	0.549**	-0.256	-0.147	0.462**	0.122	0.169	1			
10 积极面孔相似的陌生面孔吸引力评分	0.315*	-0.067	-0.415**	0.567**	0.218	-0.054	0.511**	0.096	-0.114	1		
11 中性面孔相似的陌生面孔吸引力评分	0.058	0.377*	-0.209	0.117	0.359*	-0.194	0.134	0.443**	-0.07	0.329*	1	
12 消极面孔相似的陌生面孔吸引力评分	-0.161	-0.316*	0.319*	-0.142	-0.091	0.571**	-0.004	-0.16	0.524**	-0.023	-0.2	1

注: ** $p < 0.01$, * $p < 0.05$, 下同。

由表2的结果可以发现,熟悉面孔的特质评分、熟悉面孔的吸引力评分和陌生面孔的吸引力评分在积极、消极和中性三个相应水平上均有显著的相关。为了进一步考察熟悉面孔的特质评分是否通过影响其吸引力评分(发生晕轮效应),进而影响陌生面孔的吸引力评分(出现泛化效应),本研究采用了中介效应分析的方法进行检验。熟悉面孔的特质评价、熟悉面孔的吸引力评价和陌生面孔吸引力评价之间的相关满足中介效应检验的条件。因此,本研究采用偏差校正的 Bootstrap 法(方杰,张敏强,邱皓政,2012)检验熟悉面孔吸引力评价在熟悉面孔特质评价和陌生面孔吸引力评价之间的中介作用。

检验的结果如表3所示,95%置信区间不包含0即表示中介效应成立。积极特质熟悉面孔的特质评分通过影响其吸引力评分,进而对相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平;消极特质熟悉面孔的特质评分也通过影响其吸引力评分,进而对与其相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平;中性特质的熟悉面孔则未发生相应的影响。这表明熟悉面孔的特质评价(无论是积极还是消极)均可以通过影响其吸引力评分,进而影响与其相似的陌生面孔吸引力评分,即陌生面孔的吸引力评价是由熟悉面孔特质发生晕轮效应后,进而出现吸引力的泛化效应而产生的。

消极特质的熟悉面孔特质评分与相似陌生面孔的特质评分和相似陌生面孔吸引力评分相关显著。经中介检验发现,消极特质熟悉面孔的特质评分通过影响其特质评分,进而对与其相似的陌生面孔吸引力评分产生影响的间接效应也达到显著水平。而积极特质的熟悉面孔评分与相似陌生面孔的特质评分、中性面孔特质评分与相似陌生面孔特质评分之间的相关不

显著，不满足中介检验的条件。这表明熟悉面孔的消极特质评价可以通过影响其特质评分，进而影响与其相似的陌生面孔吸引力评分。

表 3 中介效应显著性检验的 Bootstrap 分析

路径	效应值	SE	Boot 95% CI
积极熟悉面孔特质评分—积极熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.16	0.13	[0.097, 0.667]
中性熟悉面孔特质评分—中性熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.13	0.10	[-0.001, 0.419]
消极熟悉面孔特质评分—消极熟悉面孔吸引力评分—相似的陌生面孔吸引力评分	0.28	0.10	[0.128, 0.540]
积极熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.09	0.08	[-0.042, 0.289]
中性熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.16	0.13	[-0.011, 0.491]
消极熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.30	0.15	[0.037, 0.631]

实验 2 的数据分析修改如下：

3.3.2 学习效果检验

对熟悉面孔的热情和能力评价进行单因素方差分析发现，熟悉面孔的热情评分在不同效价水平上的得分差异显著，($F(2, 36) = 53.92, p < 0.001, \eta_p^2 = 0.50$)。事后多重比较发现，与积极热情句子匹配的面孔评分显著高于与中性热情句子和消极热情句子匹配的面孔评分，与中性热情句子匹配的面孔评分显著高于与消极热情句子匹配的面孔评分($ps < 0.001$)。熟悉面孔的能力评分在不同效价水平上的得分差异也显著，($F(2, 36) = 70.19, p < 0.001, \eta_p^2 = 0.57$)。事后多重比较发现，与积极能力句子匹配的面孔评分显著高于与中性和消极能力句子匹配的面孔评分，与中性能力句子匹配的面孔评分显著高于与消极能力句子匹配的面孔评分($ps < 0.001$)。以上结果表明学习后的面孔特质获得了和行为描述句一致的效价，学习的效果显著。

3.3.3 熟悉面孔特质对陌生面孔吸引力的影响

三组陌生面孔在吸引力评分上差异显著，($F(2,36) = 4.93, p = 0.009, \eta_p^2 = 0.08$)。事后多重比较发现，与积极面孔相似的陌生面孔吸引力评分显著高于与消极面孔相似的陌生面孔吸引力，与中性面孔相似的陌生面孔吸引力评分处于两者之间，与两者的吸引力评分之间无显著差异($ps > 0.05$)。这表明，通过行为描述习得的熟悉面孔特质对相似陌生面孔吸引力评价产生影响。

3.3.4 熟悉面孔特质对其吸引力评价的影响

对熟悉面孔的吸引力进行单因素方差分析发现，与不同句子效价匹配的熟悉面孔在其吸引力评分上差异显著，($F(2, 36) = 19.96, p < 0.001, \eta_p^2 = 0.27$)。事后多重比较发现，与积极描述句子匹配的面孔吸引力评分显著高于与中性句子和消极句子匹配的面孔吸引力评分，与中性句子匹配的面孔吸引力评分显著高于与消极句子匹配的面孔吸引力评分($ps < 0.05$)。这表

明, 通过行为句子描述而形成的熟悉面孔特质对其吸引力的评分产生了影响, 与积极特质匹配的面孔相比于消极特质匹配的面孔有更高的吸引力, 即熟悉面孔的特质发生了晕轮效应。

3.3.5 熟悉面孔特质对陌生面孔特质的影响

对陌生面孔的热情、能力评分进行单因素方差分析发现, 陌生面孔在热情评分上差异显著, ($F(2, 36) = 9.53, p < 0.001, \eta^2_p = 0.15$)。事后多重比较发现, 与积极热情面孔相似的陌生面孔评分显著高于与消极热情面孔相似的陌生面孔评分, 与中性热情面孔相似的陌生面孔评分也显著高于与消极热情面孔相似的陌生面孔评分, 与积极热情面孔和中性热情面孔相似的陌生面孔评分之间没有显著差异。这表明, 陌生面孔的热情特质评价受到了熟悉面孔特质效价的影响, 陌生面孔的特质与熟悉面孔产生了效价一致现象, 即熟悉面孔的热情特质发生了泛化。

陌生面孔在能力评分上的差异显著, ($F(2, 36) = 16.73, p < 0.001, \eta^2_p = 0.24$)。事后多重比较发现, 与积极能力面孔相似的陌生面孔能力评分显著高于与中性能力和消极能力面孔相似的陌生面孔能力评分, 与中性能力面孔相似的陌生面孔能力评分显著高于与消极能力面孔相似的陌生面孔能力评分($ps < 0.05$)。这表明, 陌生面孔的能力评价也受到了熟悉面孔特质效价的影响, 陌生面孔的能力特质与熟悉面孔也产生效价一致现象, 即熟悉面孔能力特质发生了泛化。

3.3.6 熟悉面孔特质对陌生面孔吸引力影响的路径分析

表 5 各变量相关分析结果

变量	1	2	3	4	5	6	7	8	9	10	11	12
1 积极熟悉面孔特质评分	1											
2 中性熟悉面孔特质评分	0.272	1										
3 消极熟悉面孔特质评分	-0.506**	-0.020	1									
4 积极熟悉面孔吸引力评分	0.573**	0.175	-0.273	1								
5 中性熟悉面孔吸引力评分	0.288	0.600**	-0.087	0.218	1							
6 消极熟悉面孔吸引力评分	-0.318	0.076	0.395*	-0.219	-0.084	1						
7 积极面孔相似的陌生面孔特质评分	0.477**	0.385*	-0.026	0.374*	0.313	-0.076	1					
8 中性面孔相似的陌生面孔特质评分	-0.061	0.408*	0.248	0.091	0.291	-0.016	0.165	1				
9 消极面孔相似的陌生面孔特质评分	-0.239	-0.065	0.540**	-0.072	-0.114	0.465**	-0.107	0.171	1			
10 积极面孔相似的陌生面孔吸引力评分	0.496**	0.155	-0.148	0.587**	0.2	-0.188	0.598**	0.065	0.037	1		
11 中性面孔相似的陌生面孔吸引力评分	-0.088	0.399*	0.301	0.16	0.371*	0.275	0.391*	0.465**	0.422**	0.431**	1	
12 消极面孔相似的陌生面孔吸引力评分	-0.134	0.085	0.481**	-0.103	0.177	0.618**	0.077	0.268	0.687**	0.096	0.546**	1

由表 5 的结果可知, 熟悉面孔的特质评分、熟悉面孔的吸引力评分和相似的陌生面孔吸引力评分之间在积极、消极、中性三个对应的水平上均达到显著相关; 熟悉面孔的特质评分、

相似陌生面孔的特质评分和相似陌生面孔的吸引力评分之间在积极、消极、中性三个对应的水平上也达到了显著的相关，均满足进行中介效应检验的基本条件。

进一步本研究采用了中介效应分析方法，检验熟悉面孔的特质是通过何种途径对陌生面孔吸引力发生作用的。我们通过偏差矫正的 Bootstrap 法考察了熟悉面孔的吸引力评分在熟悉面孔特质评分和相似的陌生面孔吸引力评分之间的中介作用，以及相似陌生面孔的特质评分在熟悉面孔特质评分和相似陌生面孔的吸引力评分之间的中介作用，分析结果见表 6。

积极特质熟悉面孔的特质评分通过影响其吸引力评分，进而对相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平；同时，消极特质熟悉面孔的特质评分通过影响其吸引力评分进而对相似的陌生面孔吸引力评分产生影响的间接效应也达到显著水平。该结果表明，熟悉面孔的特质可以通过影响其吸引力评分，进而影响陌生面孔的吸引力评分，即熟悉面孔的特质首先发生晕轮效应，进而通过泛化效应影响陌生面孔的吸引力加工。

积极特质熟悉面孔的特质评分也可以影响与其相似陌生面孔的特质评分，进而对相似的陌生面孔吸引力评分产生影响的间接效应达到显著水平；同时，熟悉面孔的消极特质评分影响与其相似的陌生面孔的特质评分，进而对相似的陌生面孔吸引力评分产生影响的间接效应也达到显著水平。该结果表明，熟悉面孔的特质评分可以通过影响陌生面孔的特质评分，进而影响陌生面孔的吸引力评分，即熟悉面孔的特质可以先通过泛化，进而再发生晕轮效应影响陌生面孔的吸引力加工。

表 6 中介效应显著性检验的 Bootstrap 分析

路径	效应值	SE	Boot 95% CI
积极熟悉面孔的特质评分—积极熟悉面孔的吸引力评分—相似的陌生面孔吸引力评分	0.23	0.10	[0.056, 0.434]
中性熟悉面孔的特质评分—中性熟悉面孔的吸引力评分—中性的陌生面孔吸引力评分	0.14	0.16	[-0.146, 0.479]
消极熟悉面孔的特质评分—消极熟悉面孔的吸引力评分—相似的陌生面孔吸引力评分	0.22	0.13	[0.038, 0.560]
积极特质熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.20	0.89	[0.068, 0.422]
中性特质熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.16	0.13	[-0.111, 0.491]
消极特质熟悉面孔特质评分—相似的陌生面孔特质评分—相似的陌生面孔吸引力评分	0.37	0.14	[0.150, 0.720]

具体修改之处见本文档 P29“2.3.2”至 P30“2.3.6”以及 P34“3.3.2”至 P35“3.3.6”标蓝部分。

数据分析中如仍有不足之处请专家继续批评指正。

意见 7：在关于遗传与环境的讨论中，作者的研究结果只能表明后天的经验存在作用，并不能否定遗传的作用，因此不能说“对于这一争论,本研究的结果也支持了后者的观点”，请注意结论的严谨性！不要为了增加研究价值，妄下结论！作者整篇文章中，部分结论存在问

题，请仔细思考。

回应：感谢审稿专家的意见以及严谨的治学态度。诚如审稿专家所言，我们的研究只能表示后天的经验存在作用，而不能否认遗传的作用；“对于这一争论，本研究的结果也支持了后者的观点”的表述欠妥。我们经过仔细思考，确实发现本研究并不能说明是遗传还是环境的作用更大的问题，只能说明环境会对偏好产生影响。因此，我们对后面关于遗传和环境讨论部分进行了修改，如“遗传和后天环境都影响着个体对美貌的偏好”，“对于高水平的知觉和认知过程遗传所发挥的作用更大”改为“这说明遗传对高水平的面孔知觉和认知过程有重要的影响”而本文的研究结果强调的是学习经验的泛化，即后天环境的作用，最后我们表述为“因此，在面孔的审美过程中后天的经验也起着重要作用。”**具体修改之处见本文档 P39 最后一段字体标蓝部分**。如果讨论仍存在不严谨之处，请专家继续批评指正。

第二轮

审稿人 1

意见 1：作者的回复解决了大部分我提出的问题。但是有一个核心的问题可能需要作者进一步的阐释。作者主要的结果都是建立在学习对所谓“陌生面孔”的影响之上，其关键在于所谓的“陌生面孔”到底在多大程度上是“陌生”的。作者认为，由 50% 熟悉面孔和 50% 陌生面孔合成的面孔就足以让被试认为这个面孔是陌生面孔。但是 50% 只是一个物理上的概念，实际上被试是否认为这是个陌生面孔，我们不得而知。万一被试根据某些特征，认为 50%A50%B 的面孔是自己见过的熟悉面孔，则作者发现的效应是否就可以由记忆而不是泛化所解释了？所以我建议作者再补充一个测试，说明被试确实觉得 50% 的熟悉面孔是未曾见过的新面孔，或者被试对这些面孔的熟悉感到底有多大。

回应：非常感谢审稿专家的意见，我们进行如下回应：根据 Verosky 和 Todorov(2010)的研究，考察面孔相似性和面孔“新/旧”主观判断关系的方法：首先在屏幕中间呈现一个 500ms 的注视点“+”，然后屏幕中间呈现 50%A50%B 合成面孔，左右两边分别呈现 100%A 熟悉面孔和 100%B 陌生面孔，要求被试对中间 50%A50%B 的合成面孔进行分类，判断中间的合成面孔与左边的面孔是同一个人还是与右边的面孔是同一个人。

在补充测试 1 中，我们首先采用 Verosky 和 Todorov(2010)的方法，共招募被试 34 人，其中男 15 人、女 19 人，年龄在 19~22 岁之间($M = 20.03$, $SD = 1.74$)。结果发现，被试将相似

度为 50%的合成面孔判断为完全陌生面孔的比例($M = 59%$)显著高于判断为熟悉面孔的比例($M = 41%$), $\chi^2 = 12.05$, $df = 1$, $p < 0.001$, 具体测试结果见下文表 1。本次实验的结果与 Verosky 和 Todorov(2010)的结果一致, 即当合成面孔相似度为 50%时, 被试更倾向于将这张合成面孔判断为是一张陌生的新面孔, 而不是更倾向于将其判断为是原来的那张熟悉面孔, 这说明 50%A50%B 的合成面孔已经足以让被试认为这是一张新的陌生面孔。

另外, 我们又根据专家的意见进一步做了测验 2, 让被试对 50%A50%B 合成面孔的熟悉度进行评价。测验 2 共招募被试 30 人, 男 14 人, 女 16 人, 年龄在 18~24 岁之间($M = 19.20$, $SD = 2.04$)。首先我们要求被试对面孔与特质词汇进行匹配学习(与实验一学习阶段的程序一致), 在评价阶段要求被试对学习过的熟悉面孔 A、50%A50%B 合成面孔和完全陌生面孔的熟悉度进行 7 点评分(1 = 完全不熟悉, 2 = 比较不熟悉, 3 = 不太熟悉, 4 = 不确定, 5 = 有点熟悉, 6 = 比较熟悉, 7 = 非常熟悉)。结果发现, 被试对三类面孔的熟悉度差异显著, $F(2, 29) = 124.25$, $p < 0.001$, $\eta^2_p = 0.49$ 。其中, 被试对熟悉面孔 A 的熟悉度评分($M = 5.72$, $SD = 0.82$)显著大于 50%A50%B 合成面孔的熟悉度($M = 3.77$, $SD = 1.02$), 显著大于陌生面孔的熟悉度($M = 1.93$, $SD = 0.94$), 而 50%A50%B 合成面孔的熟悉度显著大于完全陌生面孔的熟悉度($ps < 0.01$), 具体的测试结果见下文表 2。结果表明, 被试对 50%合成面孔的熟悉度与其他两类面孔的熟悉度存在显著差异, 说明 50%A50%B 合成面孔的效应并不是熟悉度带来的影响。

第三, 在其他考察面孔泛化效应的研究中, 研究者们也大多采用 50%的合成面孔作为新的面孔材料, 发现面孔相似度为 50%的情况下, 其泛化效应较强, 因此许多研究者使用 50%的相似度作为考察泛化效应的条件(DeBruine, 2005; Verosky, Todorov, 2012; Günaydin, Zayas, Selcuk, & Hazan, 2012; Gawronski & Quinn, 2013)。基于此, 本研究根据前人的研究经验也选取了 50%的相似度作为面孔泛化效应发生的条件。

综上所述, 被试并不会将 50%的合成面孔完全判断为熟悉面孔, 我们认为被试在判断时依据面孔的相似性而产生与熟悉面孔一致的反应, 是对熟悉面孔产生了泛化效应。

涉及的相关文献:

DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proceedings Biological Sciences*, 272(1566), 919-922.

Gawronski, B., & Quinn, K. A. (2013). Guilty by mere similarity: assimilative effects of facial resemblance on automatic evaluation. *Journal of Experimental Social Psychology*, 49(1), 120-125.

Günaydin, G., Zayas, V., Selcuk, E., & Hazan, C. (2012). I like you but i don't know why: objective facial

resemblance to significant others influences snap judgments. *Journal of Experimental Social Psychology*, 48(1), 350-353.

Verosky, S. C., & Todorov, A. (2010). Generalization of affective learning about faces to perceptually similar faces. *Psychological Science*, 21(6), 779-785.

Verosky, S. C., & Todorov, A. (2013). When physical similarity matters: mechanisms underlying affective learning generalization to the evaluation of novel faces. *Journal of Experimental Social Psychology*, 49(4), 661-669.

测试结果如下表:

表 1 测试 1 被试对 50%合成面孔判断的结果

被试编号	判断为陌生面孔的比例	判断为熟悉面孔的比例
1	0.4	0.6
2	0.8	0.2
3	0.3	0.7
4	0.5	0.5
5	0.5	0.5
6	0.6	0.4
7	0.7	0.3
8	0.7	0.3
9	0.5	0.5
10	0.8	0.2
11	0.6	0.4
12	0.7	0.3
13	0.9	0.1
14	0.6	0.4
15	0.4	0.6
16	0.6	0.4
17	0.5	0.5
18	0.6	0.4
19	0.5	0.5
20	0.4	0.6
21	0.7	0.3
22	0.5	0.5
23	0.7	0.3
24	0.4	0.6
25	0.6	0.4
26	0.6	0.4
27	0.6	0.4
28	0.5	0.5
29	0.7	0.3
30	0.8	0.2

31	0.7	0.3
32	0.7	0.3
33	0.5	0.5
34	0.6	0.4

表 2 测试 2 被试对三类面孔熟悉度的判断结果

被试编号	面孔熟悉度评分		
	学习过的熟悉面孔	50%的合成面孔	完全陌生面孔
1	6.17	3.17	2.33
2	5.33	5.50	4.50
3	7.00	4.17	1.00
4	5.83	2.67	1.00
5	4.50	4.00	1.50
6	4.50	2.50	2.00
7	7.00	2.50	1.00
8	5.83	2.33	1.00
9	5.00	5.33	2.50
10	5.17	2.17	3.00
11	4.33	2.67	2.00
12	7.00	4.17	1.00
13	5.17	4.00	3.00
14	5.67	3.83	2.50
15	7.00	4.17	1.00
16	5.67	4.17	3.83
17	5.33	3.17	2.17
18	5.00	3.83	1.00
19	6.00	2.83	1.00
20	6.67	4.50	1.00
21	7.00	5.83	2.33
22	6.33	4.00	2.17
23	5.67	4.67	2.50
24	5.33	2.00	1.00
25	6.00	3.50	2.83
26	6.00	4.50	2.33
27	4.33	3.00	1.00
28	5.67	4.83	2.33
29	6.00	4.67	1.00
30	5.00	4.33	2.00

审稿人 2

意见 1: 作者对审稿专家的意见均作出了较好的回答, 数据分析方法合理, 结论得当, 建议发表。

回应：感谢您的肯定！

第三轮

主编终审

作者较好地回答了前两轮审稿人的建议。文章的逻辑和设计没有什么问题，但是写作仍需进一步完善。建议如下：

意见 1： P28，作者修改后的标题在内容覆盖了整个研究，但是在意义表达上感觉略微有些不足。本研究是探讨个体对陌生面孔进行较高吸引力评价的原因，而不是对陌生面孔吸引力的加工。建议作者再考虑一下。

回应：非常感谢审稿专家的宝贵意见，我们作如下回应：在面孔吸引力的研究中，面孔自身因素和观察者因素对面孔吸引力的影响是研究者关注的焦点问题。一些研究者认为面孔吸引力是面孔自身的特点所决定，即“beauty is only skin deep”(Zangwill, 2000; Rhodes & Tremewan, 2010; Vingilis-Jaremko & Maurer, 2013)；而另一些研究者则认为面孔吸引力存在于观察者的知觉空间之中，即“beauty is in the eye of the beholder”(Little, Jones, Penton-Voak, Burt, & Perrett, 2002; Kwart, Foulsham, & Kingstone, 2012; Thiruchselvam, Harper, & Homer, 2016)。本研究更加强调观察者因素在面孔吸引力评价中的作用，即“beauty is in the eye of the beholder”。

一方面，本研究探讨了陌生面孔吸引力评价较高的原因：熟悉面孔与特质的匹配学习，使个体对陌生面孔的吸引力评价发生了变化，表明观察者所具有的面孔相关经验对陌生面孔吸引力的评价有重要影响；另一方面，本研究非仅仅考察面孔相关经验对陌生面孔吸引力评价的影响，这种影响是如何发生的（陌生面孔吸引力的加工过程）我们也进行了探讨。通过中介分析发现，熟悉面孔的特质经验通过熟悉面孔的吸引力和陌生面孔的特质经验影响了陌生面孔的吸引力。因此，本文回答了熟悉面孔相关经验是否影响陌生面孔吸引力，以及通过什么（怎样起作用）影响了陌生面孔吸引力，即陌生面孔吸引力评价差异的来源和加工过程两个关键问题。

通过中介作用分析加工过程的方法在面孔吸引力的研究中已经得到运用，例如在 2017 年在 *Journal of Personality and Social Psychology* 发表的一篇文章中，Carr 等人(2017)探讨了曝光过程中原型面孔与合成面孔吸引力之间的差异，通过记忆模型来解释熟悉度在其中的中介作用，阐释了记忆加工过程对面孔偏好的影响。

综上，本研究即探讨了个体对陌生面孔吸引力评价较高的原因，也探讨了陌生面孔吸引力的加工过程。

结合审稿专家的建议以及我们课题组的反复讨论和斟酌，我们将中文题目改为“**美在观察者眼中：陌生面孔吸引力评价中的晕轮效应和泛化效应**”，同时将英文题目改为“**Beauty is in the eye of the beholder: The halo effect and generalization effect in the facial attractiveness evaluation**”，见修改稿第 32 页和第 51 页标红的中英文标题。

涉及的相关文献：

Carr, E. W., Huber, D. E., Pecher, D., Zeelenberg, R., Halberstadt, J., & Winkielman, P. (2017). The Ugliness-in-averageness effect: Tempering the warm glow of familiarity. *Journal of Personality and Social Psychology, 112*(6), 787-812.

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 269*(1496), 1095–1100.

Kwart, D. G., Foulsham, T., & Kingstone, A. (2012). Age and beauty are in the eye of the beholder. *Perception, 41*(8), 925-938.

Rhodes, G., & Tremewan, T. (2010). Averageness, exaggeration, and facial attractiveness. *Psychological Science, 7*(2), 105-110.

Thiruchselvam, R., Harper, J., & Homer, A. L. (2016). Beauty is in the belief of the beholder: Cognitive influences on the neural response to facial attractiveness. *Social Cognitive and Affective Neuroscience, 11*(12), 1999–2008.

Vingilis-Jaremko, L., & Maurer, D. (2013). The influence of symmetry on children's judgments of facial attractiveness. *Perception, 42*(3), 302-320.

Zangwill, N. (2000). Skin deep or in the eye of the beholder? the metaphysics of aesthetic and sensory properties. *Philosophy & Phenomenological Research, 61*(3), 595-618.

意见 2： P33 2.2.2 实验材料部分和 P38 3.2.2 实验材料部分所使用的词汇和句子，作者是否可以在文中进行示例展现？

回应： 非常感谢审稿专家的意见，我们按照审稿专家的要求将实验材料部分所使用部分词汇和句子添加在正文中。由于句子材料内容较多，我们在 3.2.2 实验材料部分进行了示例展现。

见修改稿第 36 页“2.2.2 实验材料”和第 41 页“3.2.2 实验材料”标红部分。

意见 3: 写作规范，如统计量斜写、中英文括号混用等.....

回应: 非常感谢审稿专家的宝贵意见。我们根据《心理学报》最新的样刊，对文中统计量的书写进行了检查，没有用斜体书写的统计量已经纠正过来，所有的统计量、符号和数字之间增加了空格，所有的括号统一为英文括号，文中修改之处较多，在此不便一一指出。如仍有书写不当之处，请专家继续批评指正并给予我们继续修改的机会。